Artificial music theory:
Interpreting the hidden spaces in neural-network models as a tool for music-theoretical investigation

Gabriele Cecchetti*, Roger T. Dean
The MARCS Institute for Brain, Behaviour, and Development (Western Sydney University)
*Gabriele.Cecchetti@westernsydney.edu.au

A trained generative neural network (NN) encodes structural knowledge about the corpus it is trained on (Manning et al., 2020): yet, while AI-based music generation attracts increasing interest (Ji et al., 2023), the potential of interpretable NNs as tools for music-theoretical enquiry remains under-exploited. Here, we investigate the geometry of a NN's hidden embeddings as representative of stylistic variation across pieces, as well as structural features within pieces, with the qualitative exploration of individual examples and quantitative hypothesis tests over large corpora.

An LSTM-based NN was trained for next-note prediction on annotated, score-based (Hentschel, 2024) and non-annotated, performance-based corpora (Hawthorne et al., 2019) of tonal repertoires. We show that the LSTM hidden states encode aspects of musically meaningful features such as voice segregation, chord identity, and harmonic directedness. This learning is theory-agnostic (the model does not commit to any music-theoretical notions), bottom-up (the model is only exposed to a raw MIDI representation of the musical surface), and unsupervised (the LSTM states have no explicit reward for encoding features of the individual events). This suggests that meaningful structural principles plausibly emerge spontaneously in the service of successful prediction. In turn, this provides a useful common ground for competence-level theories of musical structure (e.g., Lerdahl & Jackendoff, 1983) and cognitive accounts based on statistical learning (Pearce, 2018) and predictive-coding principles (Koelsch et al., 2019).

We finally discuss applications of these tools to a corpus of free (non-idiomatic; Bailey, 1980) improvisations (Dean & Forth, 2020), whose aesthetics resists explicit a-priori theorizing (Canonne, 2018), and where structural principles determined a priori coexist with those emerging in-the-moment during performance (Dean & Bailes, 2016; Sparti, 2016) or cemented implicitly over repeated improvisations (Smith & Dean, 1997). In these repertoires, interpretable NNs may prove to be a valuable analytical tool to identify and creatively exploit data-driven yet abstract structural and stylistic features with minimal a-priori theoretical commitment.

Presentation format: Oral presentation

REFERENCES

Bailey, D. (1980). *Improvisation: Its Nature And Practice In Music*. Moorland publishing.

Canonne, C. (2018). Rehearsing Free Improvisation? An Ethnographic Study of Free Improvisers at Work. *Music Theory Online*, *24*(4). https://mtosmt.org/issues/mto.18.24.4/mto.18.24.4.canonne.php

Dean, R. T., & Bailes, F. (2014). Cognitive Processes in Musical Improvisation. In G. E. Lewis & B. Piekut (Eds.), *The Oxford Handbook of Critical Improvisation Studies* (Vol. 1). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780195370935.013.007

Dean, R. T., & Forth, J. (2020). Towards a Deep Improviser: A Prototype Deep Learning Post-Tonal Free Music Generator. *Neural Computing and Applications*, *32*, 969–979.

Hawthorne, C., Stasyuk, A., Roberts, A., Simon, I., Huang, C.-Z. A., Dieleman, S., Elsen, E., Engel, J., & Eck, D. (2019). *Enabling Factorized Piano Music Modeling and Generation with the MAESTRO Dataset* (arXiv:1810.12247). arXiv. https://doi.org/10.48550/arXiv.1810.12247

Hentschel, J. (2024). *Data-Driven Music Theory: Curating and Investigating Large Corpora of Digitally Encoded Music Analyses* [EPFL]. https://doi.org/10.5075/epfl-thesis-10276

Ji, S., Yang, X., & Luo, J. (2023). A Survey on Deep Learning for Symbolic Music Generation: Representations, Algorithms, Evaluations, and Challenges. *ACM Comput. Surv.*, *56*(1), 7:1-7:39. https://doi.org/10.1145/3597493

Koelsch, S., Vuust, P., & Friston, K. (2019). Predictive Processes and the Peculiar Case of Music. *Trends in Cognitive Sciences*, *23*(1), 63–77. https://doi.org/10.1016/j.tics.2018.10.006

Lerdahl, F., & Jackendoff, R. S. (1983). *A Generative Theory of Tonal Music*. MIT Press.

Manning, C. D., Clark, K., Hewitt, J., Khandelwal, U., & Levy, O. (2020). Emergent Linguistic Structure in Artificial Neural Networks Trained by Self-Supervision. *Proceedings of the National Academy of Sciences*, *117*(48), 30046–30054. https://doi.org/10.1073/pnas.1907367117

Pearce, M. T. (2018). Statistical Learning and Probabilistic Prediction in Music Cognition: Mechanisms of Stylistic Enculturation. *Annals of the New York Academy of Sciences*, *1423*(1), 378–395. https://doi.org/10.1111/nyas.13654

Smith, H., & Dean, R. T. (1997). *Improvisation Hypermedia and the Arts since 1945*. Routledge. https://doi.org/10.4324/9781315077994

Sparti, D. (2016). On the Edge: A Frame of Analysis for Improvisation. In G. E. Lewis & B. Piekut (Eds.), *The Oxford Handbook of Critical Improvisation Studies* (Vol. 1). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780195370935.013.020